# Learning the Lessons of Tit-for-Tat: Even Competitors Can Get the Message

Kennon M. Sheldon
University of Missouri—Columbia

The tit-for-tat strategy is known for its ability to train cooperation during an iterated 2-person prisoner's dilemma. In this research, a repeated measures design was used to assess the nature and long-term persistence of such training. One month after playing against a programmed tit-for-tat strategy, participants returned to the lab to play a randomly matched other participant. Participants increased their levels of cooperative behavior, both during the 1st sequence and in the 2nd sequence compared with the first. Dispositional competitors increased the most from the 1st to the 2nd sequence and also gained the most new respect for opponents' self-interestedness following the 1st sequence. This suggests that competitors may be individualists at heart and that the "punitive" feature of tit-for-tat (R. Axelrod, 1984) is important for eliciting enhanced cooperation from such opponents.



The world needs cooperation, now more than ever. In the next century, humans will become increasingly interdependent (Rusbult & Van Lange, 1996) as they struggle to cope with the increasing complexity of society and with diminishing resources in the world (Axelrod, 1997; Laszlo, Artigiani, Combs, & Csanyi, 1996). Only by ever more coordinated functioning will humans be able to overcome these challenges (Sober & Wilson, 1998). How can lasting cooperation be achieved?

A large volume of research has examined factors that can help train cooperation within the immediate situation. Such factors include reward structure, whether communication is allowed, the opponent's strategy, and induced positive mood (see Komorita & Parks, 1995, for a review). Despite the successes of this research, it is surprising that researchers have not examined whether training lasts—after having engaged in a social dilemma under cooperation-enhancing conditions, is one thus more likely to cooperate in the future? Also, are some personality types better prepared to learn cooperative lessons than others?

I explored these questions by asking participants to participate in two prisoner's dilemma (PD) sequences, held 1 month apart. In the first sequence, participants played against a programmed tit-for-tat strategy (which is often successful at training immediate cooperation in its opponents; Oskamp, 1971), and in the second sequence participants played against randomly matched other participants. Participants' expectancies regarding their opponent were also assessed twice prior to each sequence. This repeated measures design allowed examination of the stability of any behavioral changes induced by tit-for-tat, and it also allowed assessment of within-subject changes in expectancies, both as a function of having faced tit-for-tat and as predictors of subsequent changes in cooperative behavior. In addition, I incorporated a between-subjects variable into the study, namely, participants' social value orientation (Liebrand, 1984; Messick & McClintock, 1968). This allowed examination of the effects of personality dispositions on behavior and expectancy outcomes. That is, to what extent do dispositional cooperators, individualists, and competitors differ in revising their expectancies and behavior after facing tit-for-tat? Background issues are considered below.

## Defining the Cooperative Act

The act of cooperation involves a state of hopeful vulnerability in which one makes an offer while simultaneously incurring a risk. In terms of the classic PD, the risk occurs because of the fact that one will score the least amount of points if one's offer is not simultaneously reciprocated by the other. In fact, it is most rational in a single play or short-term situation to assume that the other will not reciprocate, and thus to defect in self-defense. Another advantage of defection is that one accrues the most amount of points if the other cooperates on that trial. Despite the strategic dominance of defection (Van Lange, Liebrand, Messick, & Wilke, 1992), cooperation becomes the most rational regime in the long term if the parties can arrive at and maintain it (Axelrod, 1984). This is because sustained cooperation provides the best mutual or "pareto-optimal" solution to the PTD (Komorita & Parks, 1995) in that both individuals score more if they both cooperate than if they both defect. However, it is never certain that sustained cooperation can be achieved, because outcomes are not controllable by any one individual.

Correspondence concerning this article should be addressed to Kennon M. Sheldon, Department of Psychology, University of Missouri—Columbia, 210 McAlester Hall, Columbia, Missouri 65211. Electronic mail may be sent to SheldonK@missouri.edu.

Journal of Personality and Social Psychology, 1999, Vol. 77, No. 6, 1245–1253
 0022-3514/99/$3.00

## How Is Sustained Cooperation Achieved? The Goal-Expectation Theory

Pruitt and Kimmel's (1977) goal-expectation theory says that people will cooperate only when they both have the goal of achieving mutual cooperation and the expectation that the other will cooperate in return. Pruitt and Kimmel argued that the situational goal of attaining cooperation typically arises only after the participant learns that (a) exploitation is hopeless because the other will not tolerate it, (b) his or her own outcomes are dependent on the other's choices, and (c) he or she must cooperate if the opponent is to do so. Notably, Pruitt and Kimmel assumed that all persons are ultimately individualists; that is, they are concerned for maximizing their own gain within the limits set by their opponents. Although Pruitt and Kimmel's model is more than two decades old, it remains influential (De Dreu, Giebels, & Van de Vliet, 1998; Fry, Betz, & Pruitt, 1996).

## Individual Differences in Cooperative Goals

Pruitt and Kimmel (1977) focused primarily on situational factors influencing immediate cooperative goals. However, in this research I applied their model at an individual-differences level of analysis by assessing participants' general motives regarding prosocial behavior. In fact, much research indicates that the impulse to cooperate is a fundamental and stable dimension of personality on which individuals vary (Graziano & Eisenberg, 1997); constructs such as communality, agreeableness, and sociability appear within nearly every personality system. In social dilemma research, the most widely researched construct of this type is social value orientation (SVO), in which individuals are typed as *cooperators, individualists,* or *competitors* (Messick & McClintock, 1968) according to their choices within a series of decomposed prisoner's dilemma games. Persons of these different value types are assumed to make a transformation of utilities for a given payoff matrix (Kelley & Thibaut, 1978; Van Lange, 1992), so that what would be the most desirable outcome from a purely individualistic perspective (e.g., getting the most possible points for self) is not necessarily what people desire. Instead, cooperators try to maximize joint gain (by getting the most possible points for the dyad), whereas competitors try to maximize relative gain (by getting the largest possible margin of victory).

One implication of the SVO research is that Pruitt and Kimmel's (1977) "ultimate individualism" assumption may not hold true for all persons. For example, competitors may refuse to accommodate their behavior to the cooperative but punitive opponent (Van Lange et al., 1992) because a tie cannot satisfy their real motive of outdoing the other. By continuing to defect (perhaps in spite) even when this strategy has proved hopeless, competitors may fail to score as many points for themselves as they could, thereby also unduly limiting their partners. In fact, Kuhlman and Marshallo (1975) reported just such a pattern; they showed that competitors refused to accommodate to the (nonexploitable) tit-for-tat strategy; that is, competitors continued to defect throughout the PD sequence. It is clear that such a tendency may have sobering implications for human beings' potential to function together in more coordinated and harmonious ways.

## Individual Differences in Cooperative Expectations

Pruitt and Kimmel's (1977) model also implies that individual differences in cooperative expectations are important. Obviously, a person may elect not to cooperate, despite having a cooperative goal, if that person does not believe that others will reciprocate (Yamagishi, 1992). In the past, researchers have suggested that each SVO type expects others to have the same goal that they do; for example, cooperators expect others to try to maximize joint gain, individualists expect others to try to maximize their own gain, and competitors expect others to try to maximize the difference between self and other. This has been called the *structured assumed similarity bias* (SASB; Kuhlman, Brown, & Teta, 1992).

Although competitors expect others to have the same goal as they do (i.e., to achieve dominance), their exploitative behavior suggests that they may also view their opponents as "less tough on defense." That is, because competitors typically encounter persons who are less aggressive than themselves, they may learn that their opportunistic behavior often goes unpunished. Consistent with this view, several investigators have found that, in general, people are slow to punish defectors (Pruitt, 1970; Tognoli, 1975). It seems logical that competitors, who seek to maximize relative gain, are well aware of this niche of opportunity. The foregoing reasoning suggests a perhaps paradoxical pattern in which competitors believe that others will be pursuing competitive advantage (like themselves) but also believe that others will be somewhat tolerant of competitors' exploitative efforts.

To test this idea, I assessed participants' expectancies in a somewhat different way than in past research. Specifically, I measured (a) the extent to which the opponent was expected to consider the participant's interests, (b) the extent to which the opponent was expected to consider his or her own interests, and (c) the extent to which the opponent was expected to consider both players' joint interests. In line with the SASB demonstrated by past researchers, I expected that competitors would expect less cooperation from others; that is, they would not expect their opponent to consider the joint interests of both players. However, in line with the idea that competitors are accustomed to finding that they can exploit people, I hypothesized that competitors would also not expect others to be as vigilant regarding their own interests.

## Training Cooperation

Given the aforementioned variations in individuals' cooperative goals and expectancies, the task of training protagonists to cooperate with one another is clearly a daunting one. How can it be accomplished? Ample past research has shown that the tit-for-tat strategy tends to be effective at eliciting cooperation from its opponents (Komorita, Parks, & Hulbert, 1992; Oskamp, 1971). Tit-for-tat cooperates on the first move and imitates its opponent thereafter. Three features of this strategy have been shown to be particularly conducive to the elicitation of cooperation (Axelrod, 1984). First, the strategy is "nice" (it will never initiate a cycle of defection). Second, the strategy is "punitive" (as noted above, tit-for-tat instantly retaliates against defection). Finally, the strategy is "forgiving" (when the opponent returns to cooperation, tit-for-tat does so also). Notably, the punitive aspect of tit-for-tat may, in the long run, be quite relevant for training competitors to cooperate, because, as suggested above, they may be prone to believe that others are "weak on defense."

To examine the processes by which individuals may learn to take on a more cooperative approach in social dilemmas, I pitted participants against a programmed tit-for-tat strategy in a first PD sequence. Thus, all participants faced the "same opponent," strategically speaking, although of course they experienced many different game trajectories according to their own initial choices and ongoing reactions to the experience. Facing tit-for-tat was expected to give rise to two effects. First, participants should move toward increasing cooperation within the first PD sequence, such that they would make more cooperative responses at the end of the sequence compared with at the beginning of the sequence. This would concur with past findings regarding the general efficacy of tit-for-tat in training immediate cooperation (Oskamp, 1971). Second, participants should be more likely to pursue a cooperative strategy in a second PD sequence against a random other participant, 1 month later.[1] That is, I made the optimistic assumption that facing tit-for-tat has longer term, as well as short-term effects, on increased cooperation.

How might cooperators and competitors react differently after facing tit-for-tat? That is, should competitors be expected to increase their level of cooperation, or perhaps decrease it, in the second PD sequence? Two contradictory hypotheses can be offered. One prediction is that the different types, if they change, will only become more extreme. That is, perhaps only cooperators are prepared to learn tit-for-tat's lesson, giving up even their sporadic forays into defection; meanwhile, competitors may perhaps be rigid and refuse to change strategy, or perhaps defect even more to spite their competitor, thereby dooming both themselves and their opponents to obtain fewer resources than they might otherwise. In other words, competitors may refuse to accommodate to the terms dictated by tit-for-tat (Kuhlman et al., 1992; Kuhlman & Marshallo, 1975). Such a finding would not bode well for creating a more cooperative society.

A divergent prediction is that competitors will learn the lesson and manifest increased cooperation in the second PD sequence. Such behavioral accommodation would be an example of Pruitt and Kimmel's (1977) hypothesized two-stage sequence, in which the defector first learns that an exploitative strategy will not pay and then adjusts his or her behavior. This accommodation would be especially striking given that the competitors will have become more cooperative in the second sequence despite sometimes facing spontaneous defections from their opponent within that sequence (not encountered in the first sequence against tit-for-tat). Such a finding would be cause for optimism that a more balanced society can be achieved.

## Summary

The major innovation of the present research is that a repeated measures design was used to assess the long-term effects of playing against a tit-for-tat strategy, as moderated by SVO. A second innovation is that I measured participants' expectations regarding their opponents in a new way, by focusing on their expectancies regarding the opponent's level of self-interest, other interest, and joint interest. As a final innovation, SVO was measured a second time, 1 month after the second experimental session. Although personality variables are typically viewed as difficult, if not impossible, to change (Heatherton & Weinberger, 1994), I believed it worthwhile to examine whether apparent changes in dispositional competitiveness might occur as a function of the encounter with the firm but fair tit-for-tat strategy.

## Specific Hypotheses

1. Participants would move toward increasing cooperation from the beginning to the end of the first sequence against tit-for-tat, consistent with past findings concerning the general efficacy of tit-for-tat in training immediate cooperation.
2. Participants would cooperate more in the second PD sequence than in the first PD sequence, indicating stability and even generalizability of the lesson taught by tit-for-tat.
3. Dispositional competitors would show the largest increase in cooperation, compared with individualists and cooperators. This would indicate that competitors are individualists at heart, willing to change the most to get more points for themselves.
4. Participants would show increased respect for the opponent's levels of self-interest, other interest, and joint interest, after facing the firm but fair tit-for-tat strategy.
5. Dispositional competitors would show the largest increase in respect for the opponent's self-interestedness after playing against tit-for-tat. Because of the punitive feature of tit-for-tat, competitors will have learned that others are not necessarily "soft on defense."

In addition to these five hypotheses, I also explored the issue of social value change, mentioned earlier. That is, might competitors move toward a more cooperative disposition, as a result of the encounter with tit-for-tat? I did not venture predictions regarding this issue.

## Method

### *Participants and Overview*

Participants were 90 members (34 men and 56 women) of an upper-division psychology class at the University of Rochester who received extra credit for their participation. During the semester, they attended two experimental sessions. In the first experiment, participants completed a measure of SVO, rated their expectancies of the assigned opponent, and played 15 rounds of PD against a programmed tit-for-tat strategy (which they believed to be another participant). In the second experiment, 1 month later, participants again rated their expectancies of the other, then played a 15-round game against a real, randomly matched other participant. Near the end of the semester, they completed a questionnaire containing another measure of SVO, so that possible changes in SVO might be assessed.

### *First Experimental Procedure*

For the first experiment, participants were scheduled in groups of 4. After being greeted, participants were seated at individual desks lined up along one side of the room and separated by vertical partitions that prevented visual contact. First, participants completed the Kuhlman/Teta measure of SVO while listening to taped instructions based on the test authors' recommendations (Kuhlman, personal communication, January 4, 1996). Specifically, they made a series of 24 choices between two different

[1] Although pitting participants against other participants sacrificed some degree of experimental control, it also provided potentially greater external validity than is typically found in social dilemma research. To find that the lesson taught by tit-for-tat generalizes to behavior in a relatively open-ended encounter 1 month later would suggest that the lesson transfers to a wide variety of situations.

allocations of points to themselves and to another person (i.e., for the first decision, Choice A gave "self" 74 and "other" 90; Choice B gave both "self" and "other" 87). Participants were told that they would receive the sum of the points they allotted themselves and the points another (anonymous) person at the session allotted to them, and that the other person would receive the sum of the points he or she allotted him- or herself and the points the participant allotted him or her. Participants were asked to imagine that the points have value both to them and to the other person; the more accumulated by either participant, the better for that participant.

Next, participants were told that they would be matched anonymously with a different participant in attendance at the session in order to play another game involving a series of choices. A PD matrix (in which the four outcome cells were 60, 10; 10, 60; 40, 40; and 20, 20) was projected, and the scoring of the matrix and the general procedure for the game were explained via taped instructions. Participants were asked to imagine that the points had value both to themselves and to the other. After the explanation, I assessed partipants' expectancies regarding their opponent's interpersonal goals by asking them to complete a short questionnaire containing three questions. The questions were (a) "To what extent do you expect your paired person to consider their own interests?" (b) "To what extent do you expect your paired person to consider your interests?" and (c) "To what extent do you expect your paired person to consider both of your interests?" Responses were made using a 1 (*not at all*) to 7 (*very much*) scale. These three ratings constituted the "Other will consider my interests," "Other will consider his/her own interests," and "Other will consider both our interests" variables, respectively. At this point, the actual PD sequence began.

On each round, participants were asked to write a choice (A or B) on a slip of paper and to hand it to the experimenter, who then went to another room to record all 4 participants' choices. The experimenter calculated the (programmed tit-for-tat) feedback due each participant and returned an appropriately marked slip of paper to each. Participants then recorded their own and the anonymous other's choice for that round, and the number of points resulting both to themselves and to the other from this combination of choices. They then proceeded to choose for the next round. The procedure continued for 15 rounds, after which participants were dismissed. A total-cooperation variable was computed by counting cooperative choices across the 15 rounds of the sequence. In addition, variables representing the amount of cooperation in the first and last five moves of the sequence were computed.

### *Second Experimental Procedure*

The second experiment took place approximately 4 weeks after the first, using the same participants. Participants were again scheduled into groups of 4 (they were grouped with different other participants than in Experiment 1) and were seated around a square table. As in the first procedure, visual buffers were used to preclude visual contact with other participants, and participants were asked not to speak with each other. Taped instructions informed participants that they would again be playing a game with another classmate and led them through a review of the scoring of PD matrices and of the general procedure. Participants then completed a short questionnaire that contained the same three expectancy items, which they rated using the same 7-point scale. A practice trial was given to ensure that all of the participants understood the procedure and scoring, and then the procedure began. Participants played against the participant seated directly across from them (whom they could not see). After the 15th round in the sequence, participants were debriefed and released.

Participants' ratings on the three initial expectancy questions again constituted the "Other will consider my interests," "Other will consider his/her own interests," and "Other will consider both our interests" variables. Once again, a total-cooperation variable was created by summing participants' cooperative choices. As above, variables representing the amount of cooperation in the first five and the last five moves of the sequence were also computed.

For a preliminary analysis, I attempted to classify each participant as a cooperator, an individualist, or a competitor on the basis of their responses to the Kuhlman/Teta decomposed game measure. To be classifiable, participants have to show a clear preference for a particular type of decision (see Kuhlman et al., 1992, for a discussion of scoring procedures and criteria for decomposed games). Notably, only 65 of the 90 participants could be classified as cooperators ($n$ = 15), individualists ($n$ = 34), or competitors ($n$ = 16) by the SVO measure used. In fact, high nonclassification rates are quite common in research using decomposed game measures.[2] In an attempt to retain all participants for analyses, below I also present results for an alternative method of scoring such measures. In this method, the total number of points assigned to self and to other across the 24 decisions are considered as two separate, continuous variables. These variables may be said to represent participants' level of acquisitiveness for self and their level of generosity regarding others.[3] Presenting parallel results for the two different SVO scoring procedures makes it possible to assess whether the two methods yield conceptually different findings; if not, the continuous method may sometimes be preferred for reasons of statistical power.

Near the end of the semester, participants were given a take-home questionnaire packet to complete and return at the next class session. Included in this packet was the Ring Measure of Social Values (Liebrand, 1984). Like the Kuhlman/Teta measure, the Ring Measure requires participants to make 24 choices between two different allotments of points to self and other. Participants were informed that they had been anonymously paired with another participant in the study and were again asked to imagine that the points had value both for themselves and the other person. The Ring Measure is scored by summing both the points allotted to other and the points allotted to self. These two sums are treated as $x$ and $y$ coordinates that locate participants on a circle; participants are classified as cooperators, individualists, or competitors according to which portion of the circle they fall (see Liebrand, 1984, for further details). Of the 90 participants, 22 were later classified as cooperators, 54 as individualists, and 11 as competitors. Three participants, whose scores were outside the prescribed region of the circle, were unclassifiable.

## Results

### *Preliminary Analyses*

Gender was found to be unrelated to SVO, numbers of cooperative choices, and expectancies. Therefore I collapse across gender for the analyses reported below.

---

[2] For example, Kuhlman et al. (1992) reported that only 92 of 140 participants were classifiable, and Kuhlman and Marshallo (1975) reported that only 125 of 167 participants were classifiable. In an attempt to deal with the problem, I chose an alternative decomposed game methodology as a postmeasure of SVO (described later).

[3] In fact, researchers have sometimes scored decomposed game choices as continuous variables, rather than assigning participants to types (Kuhlman, Camac, & Cunha, 1986). In the present data, I found no difference between competitors and cooperators on the acquisitiveness variable ($M$s = 2,159 and 2,157, respectively), whereas individualists were significantly higher on this variable ($M$ = 2,185). Cooperators were highest in generosity ($M$ = 2,253), followed by individualists ($M$ = 2,185) and competitors ($M$ = 2,119). These patterns suggest that the ordinal sequence of SVO types primarily reflects differences in participants' willingness to give points to others, not differences in their willingness to take points for themselves.

### *Descriptive Statistics*

Table 1 contains descriptive statistics for total cooperation and the three expectancy variables for both experiments. Table 1 also gives the mean number of cooperative choices made during the first five moves and the last five moves of each of the two PD sequences.

### *Hypothesis Tests*

As can be seen in Table 1, participants cooperated more often in the last five moves ($M = 3.50$ cooperative choices) than in the first five moves ($M = 3.07$ cooperative choices) of the first PD sequence. This difference was significant, $t(89) = 2.70$, $p < .01$, consistent with my first hypothesis and with previous findings concerning the general efficacy of tit-for-tat in training immediate cooperation.[4] However, scrutiny of the data for each SVO type separately revealed a potential qualification to this pattern; although cooperators and individualists increased significantly from the first to the last part of the first sequence (both $p$s $< .05$), competitors did not. Instead, consistent with the findings of Kuhlman and Marshallo (1975), competitors made approximately the same number of cooperative choices in the last five choices of the first sequence ($M = 1.12$) as in the first five choices ($M = 1.25$). Although a repeated measures multivariate analysis of variance (MANOVA) with SVO as a between-subjects variable revealed that the interaction between SVO and trial block (first five choices and last five choices) was not significant, $F(2, 62) = 2.05$, $p < .14$, nevertheless, these data suggest that optimism regarding competitors ability to learn tit-for-tat's "lesson" may be unfounded.

To test my second and third hypotheses, I conducted a repeated measures MANOVA in which the two dependent measures were total cooperation in the first and second sequences. Both a within-subjects variable (sequence: first or second) and a between-subjects variable (SVO: competitor, individualist, or cooperator) were included in this analysis. Supporting my second hypothesis, there was a significant main effect of the sequence variable, $F(1, 62) = 4.36$, $p < .05$, indicating that cooperation increased from the first to the second sequence ($M = 10.01$ vs. $M = 10.91$; see Table 1). Notably, there was also a significant main effect of the SVO variable, $F(2, 62) = 15.07$, $p < .01$, on cooperation across the two sequences, supporting the construct validity of the SVO measure. Table 2 presents the means broken down by SVO.

Table 1
*Descriptive Statistics for Major Variables ($n = 90$)*

| Variable | *M* | *SD* |
|---|---|---|
| First sequence | | |
| "Other will consider my interests" | 3.58 | 1.69 |
| "Other will consider own interests" | 2.86 | 1.55 |
| "Other will consider both interests" | 4.51 | 1.69 |
| Cooperation in first 5 moves | 3.07 | 1.84 |
| Cooperation in last 5 moves | 3.50 | 2.01 |
| Total cooperation | 10.01 | 5.26 |
| Second sequence | | |
| "Other will consider my interests" | 4.13 | 1.57 |
| "Other will consider own interests" | 4.78 | 1.58 |
| "Other will consider both our interests" | 5.00 | 1.70 |
| Cooperation in first 5 moves | 3.62 | 1.61 |
| Cooperation in last 5 moves | 3.63 | 1.92 |
| Total cooperation | 10.91 | 5.00 |

Recall that my third hypothesis predicted an interaction between SVO and sequence, such that competitors would increase the most in their levels of cooperation from the first to the second sequence, compared with cooperators and individualists. The omnibus test of the interaction was not significant, $F(2, 62) = 1.85$, $p < .17$. However, because the conceptual focus of the hypothesis was competitors, an analysis was conducted to contrast competitors against the other two types. In this analysis the interaction was nearly significant, $F(1, 63) = 3.75$, $p = .057$. Figure 1 provides a graphic representation of this result, which suggests that there may perhaps be cause for optimism after all.

Next I turned to the three opponent expectancy variables. Three repeated measures MANOVAs were conducted; the MANOVAs included both a within-subjects variable (sequence) and a between-subjects variable (SVO). Hypothesis 4 stated that participants would evidence increases in all three expectancy constructs from the first to the second sequence. Relevant means can be found in Table 2. As can be seen, good support was found for this hypothesis: For increases in the opponent's expected concern for the participant's interests, $F(1, 62) = 2.78$, $p = .10$; for increases in the opponent's expected self-interestedness, $F(1, 62) = 33.30$, $p < .01$; and for increases in the opponent's expected joint interestedness, $F(1, 62) = 6.16$, $p < .05$. Furthermore, SVO was found to have significant main effects on the opponent's expected other-interestedness and joint-interestedness, consistent with the SASB (Kuhlman et al., 1992).

Hypothesis 5, based on Pruitt and Kimmel's (1977) model of cooperative change within social dilemmas, predicted an interaction such that competitors would show the largest increases in the expectancy that the upcoming opponent would be concerned for his or her self-interest. The omnibus test of the interaction was significant, $F(2, 62) = 4.62$, $p < .01$. Paired-sample $t$ tests revealed that both competitors, $t(15) = 4.85$, and individualists, $t(33) = 5.68$, increased in their expectancies regarding the other's self-interestedness. Figure 2 provides a graphic depiction of the results. A separate analysis was conducted to contrast competitors against the other two types, in which the SVO $\times$ Sequence interaction was also significant, $F(1, 63) = 3.92$, $p = .05$.

Next I reexamined Hypotheses 3 through 5, using the alternative scoring method for the SVO, as discussed earlier. Again, in this method two continuous variables (acquisitiveness and generosity) were created by separately summing the total points the participant gave to self and to the anonymous other during the decomposed game procedure. This method retains all participants, enhancing statistical power and the potential generalizability of findings. To find results in conceptual agreement with the type-based results would suggest that the new scoring method may confer these advantages without sacrificing information.

Table 3 presents correlations between these two variables and all cooperation and expectancy variables. Notably, the findings are essentially consistent with the mean differences by type which were reported above. It is interesting, however, that the generosity variable

[4] No such difference emerged within the second PD sequence. However, no difference was predicted because participants were not facing the cooperation-enhancing tit-for-tat strategy.

Table 2
*Mean Differences by Social Value Orientation (n = 65)*

| Type and sequence | Total cooperation | Variable: "Other will consider my interests" | Variable: "Other will consider own interests" | Variable: "Other will consider both interests" |
|---|---|---|---|---|
| First sequence | | | | |
| Cooperators | 12.87 | 4.07 | 3.60 | 5.40 |
| Individualists | 11.11 | 3.65 | 2.50 | 4.50 |
| Competitors | 4.13 | 2.75 | 2.62 | 3.69 |
| Second sequence | | | | |
| Cooperators | 13.20 | 4.60 | 4.07 | 6.07 |
| Individualists | 11.62 | 4.50 | 4.68 | 5.35 |
| Competitors | 7.13 | 2.75 | 5.75 | 4.00 |

carried the effects, whereas variations in the acquisitiveness variable were unrelated to the outcomes. Specifically, generosity was positively associated with cooperation in the first sequence but not in the second, consistent with the earlier finding that competitors became more cooperative in the second sequence (i.e., low generosity scores no longer predicted low cooperation in the second sequence).[5] Consistent with the two main effects of SVO on expectancies reported above, in both sequences generosity correlated positively with expectancies of the opponent's concern for the participant and the opponent's joint-interestedness. Consistent with the SVO × Sequence interaction on the opponent's expected self-interestedness reported above, generosity correlated positively with expectancies of the opponent's self-interestedness prior to the first sequence and negatively prior to the second sequence; in other words, those low in generosity (i.e., competitors) went from expecting relatively less self-interestedness to relatively more self-interestedness in the opponent after facing tit-for-tat.[6]

In sum, all five of the primary study hypotheses received support in these analyses. Thus, it appears that the lessons of tit-for-tat can indeed persist when participants play against each other 1 month later.[7] Finally, I turned to the issue of whether participants might actually change their social value status as a result of facing tit-for-tat. To address this question, I examined the joint distribution of the categorical variables derived from the two decomposed game measures (one measured at the beginning of the study, and one at the end). Each variable could take on a value of 1 (*cooperator*), 2 (*individualist*), 3 (*competitor*), or 4 (*unclassifiable*). Table 4 provides the cross-tabulation.

A chi-square test revealed substantial convergence between the two measures, $\chi^2(9, N = 90) = 21.36$, $p = .011$, suggesting that there is significant stability or test–retest reliability for the SVO construct. Nevertheless, there was also considerable change from the first to the second assessment of SVO. Of the 16 participants originally classified as competitors by the Kuhlman/Teta measure, only 5 were classified as competitors by the Ring Measure, and the other 11 were classified as individualists by the Ring Measure. This suggests a tendency for competitors to "reform" (Harford & Solomon, 1967) after facing tit-for-tat; that is, they will move away from competitive goals toward more individualistic goals. Furthermore, of the 34 participants classified as individualists by the Kuhlman/Teta measure, 11 were classified as cooperators by the Ring Measure and only 3 as competitors (19 were again classified as individualists by the Ring Measure, and 1 was unclassifiable). This further suggests a general learning or perhaps developmental process in which formerly self-oriented participants moved in the direction of greater prosociality after facing tit-for-tat. Although 8 of the 15 participants originally classified as cooperators by the Kuhlman/Teta measure remained cooperators according to the Ring Measure, 6 were classified as individualists by the Ring Measure (1 was unclassifiable). It is possible that some of these participants were paired with competitors during the second PD sequence, negatively affecting their joint-interestedness.

[5] Recall that the earlier predicted interaction between SVO and sequence on total cooperation did not attain significance ($p = .057$). To explicitly test this third hypothesis using the continuous method of scoring SVO, I performed a $t$ test to examine the significance of the difference between the correlation of generosity with total cooperation in the first sequence ($r = .45$) and in the second sequence ($r = .16$). The difference was significant, $t(89) = 3.03$, $p < .01$, providing unambiguous support for my third hypothesis.

[6] I also examined whether changes in perceived self-interestedness of the opponent would mediate between SVO and changes in cooperative behavior. Using the criteria for establishing mediation outlined by Judd and Kenny (1981), I conducted a regression analysis in which generosity and a difference score representing changes in the perceived self-interestedness of the opponent were entered as predictors of a difference score representing changes in cooperation. The effect of generosity was reduced (from $\beta = -.30$ to $\beta = -.22$) but not eliminated in this analysis. Thus it appears competitors' increased cooperation is not completely mediated by changes in the perceived self-interestedness of the opponent.

[7] A reviewer pointed out that potential dependencies exist within the data because participants played against other participants in the second sequence. One approach with such matched data is to use more conservative significance tests (see Griffin & Gonzalez, 1995). Adopting a .01 significance criteria for the five primary hypotheses would yield support for three out of five hypotheses, suggesting that some caution is warranted in the interpretation of results. It is notable, however, that the hypothesis tests based on the continuous method of scoring the decomposed data yielded stronger results (support for five out of five hypotheses at $p < .01$). This is expectable given the larger sample size on which these tests were based and again suggests that the typological scoring procedure may not always be the most desirable way of handling decomposed game data.

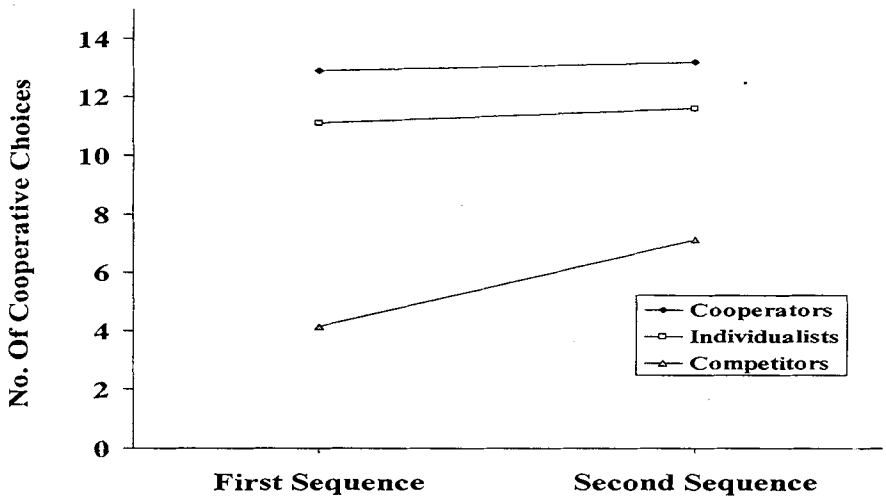


*Figure 1.* Changes in total number of cooperative choices from the first to the second sequence, as a function of social value orientation.

## General Discussion

In this research, I used a repeated measures design to examine both the immediate and longer term effects of playing against a tit-for-tat strategy in an iterated PD. This type of design has not been used in the past, and it appears that the experimental analysis of longer term change may provide important new developmental insights. Taken together, the findings have hopeful implications for the task of bringing about more harmonious and coordinated functioning within society. Participants substantially modified their behavior both within the first PD session and between the first and second sessions. In addition, participants showed flexibility in revising their expectations of their potential opponents, acquiring more "depersonalized trust" in the general goodwill of others (see, e.g., Messick & Brewer, 1983), as well as more respect for others' level of self-interestedness. Notably, these changes were demonstrated despite the fact that a full month or more had passed since the initial experience and despite the fact that in the second round some participants faced an opponent who would sometimes spontaneously defect (which tit-for-tat did not do). Thus, it appears that the "lesson" of tit-for-tat—that is, that one's opponents can be both firm and fair and that one can do well without having to "beat" the other—has substantial staying power.

Analyses involving SVO demonstrated that several of the latter findings were moderated, in meaningful and potentially important ways. Competitors, who began with the goal of maximizing their own gain relative to others, were "brought up short" by tit-for-tat

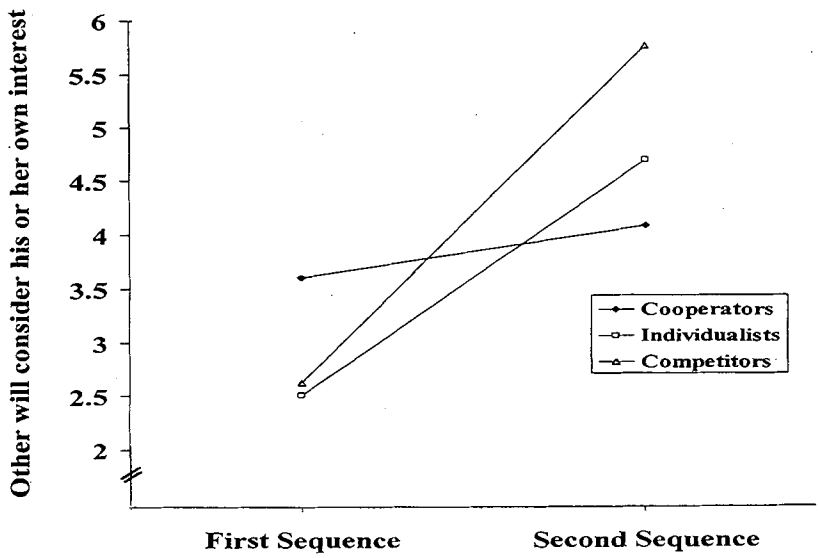


*Figure 2.* Changes in expectations of opponent's self-interestedness from the first to the second sequence, as a function of social value orientation.

Table 3
*Correlations Between the Continuously Scored Measures of Social Value Orientation and the Cooperation and Expectancy Variables (n = 90)*

| Measure | Total cooperation | Variable: "Other will consider my interests" | Variable: "Other will consider own interests" | Variable: "Other will consider both interests" |
|---|---|---|---|---|
| First sequence | | | | |
| Acquisitiveness | .04 | .02 | −.15 | .03 |
| Generosity | .45** | .24* | .23* | .29** |
| Second sequence | | | | |
| Acquisitiveness | −.01 | .08 | −.02 | .18† |
| Generosity | .16 | .28** | −.31** | .31** |

† $p < .10$ (marginally significant). * $p < .05$. ** $p < .01$.



(which will not tolerate exploitation). As a result, they were particularly likely to develop strengthened expectations regarding opponents' willingness to defend their own interests. In part because of this change, competitors increased their cooperative behavior by the largest amount from the first to the second sequence. It is interesting that this longer term change in cooperation occurred despite the fact that competitors did not increase their cooperative behavior within the first sequence against tit-for-tat, consistent with the findings of Kuhlman and Marshallo (1975). Thus, although competitors continued to defect throughout the first sequence (perhaps out of stubbornness or spite), it appears that they were willing to adopt a somewhat different strategy during the second sequence 1 month later. This can be seen as evidence of substantial behavioral accommodation having occurred, that is, the alteration of basic scripts and schemas generative of behavior in social dilemmas.

Pruitt and Kimmel's (1977) goal-expectation theory provides one framework from which to interpret the study results. Again, this theory holds that defectors learn cooperative behavior only after they discover that defection will not pay and that others can be trusted. Pruitt and Kimmel also assumed that people are all individualists at heart; that is, they are concerned with maximizing their own gain. The fact that competitors in this study evidenced a shift in their behavior after facing the nonexploitable tit-for-tat strategy, leading to their acquiring more points in the second sequence than they would have otherwise, is consistent with these two suppositions. Of course, competitors still lagged far behind the individualists and cooperators in their cooperative behavior (and in points scored). However, it may be that longer or repeated exposures to tit-for-tat might reduce the gap even further.

Pruitt and Kimmel's (1977) theory suggests that interventions to increase participants' cooperation might focus either on enhancing their expectations that others will cooperate or on their goals for cooperation. In the present study, facing tit-for-tat appears to have functioned as an intervention to alter participants' general cooperative expectations. In addition, there was some indication that competitors changed their dispositional goals as well—most of the participants who were originally typed as competitors by the Kuhlman/Teta measure of SVO were later typed as individualists by the Ring Measure, which was given after both experiments had been completed. Furthermore, many participants originally typed as individualists by the Kuhlman/Teta measure were later typed as cooperators by the Ring Measure. However, further research is needed to confirm that greater internalization of prosocial values occurs when people are exposed to the friendly, firm, and fair tit-for-tat strategy.

One final issue is worthy of discussion, namely, the alternative method for scoring decomposed game data that was explored in this data set. Because the traditional SVO typing procedure often loses many participants because of nonclassification, it seemed worthwhile to experiment with a new scoring procedure to perhaps enhance both statistical power and generalizability of results. It is reassuring that results involving this alternative procedure were conceptually similar to those involving the traditional typological scoring procedure, suggesting that no important information is lost by using the new procedure. But the results do pose a question: Why did generosity, and not acquisitiveness, account for variance in the outcomes?

Again, note that the generosity measure indexes the amount of points the participant is willing to assign an anonymous other, whereas the acquisitiveness variable indexes the amount of points the participant is willing to take for self. In the present study, cooperators and competitors did not differ from each other in their level of acquisitiveness (see Footnote 3); compared with individualists, both seemed to be willing to sacrifice individual points (cooperators in the name of achieving relative equality, competitors in the name of achieving relative dominance). However, the three SVO types did differ linearly in their levels of generosity.

Table 4
*Joint Distribution of the Kuhlman/Teta and Ring Measures of Social Value Orientation (SVO)*

| Kuhlman/Teta measure of SVO | Ring Measure of SVO: 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1. Cooperators | 8 | 6 | 0 | 1 |
| 2. Individualists | 11 | 19 | 3 | 1 |
| 3. Competitors | 0 | 11 | 5 | 0 |
| 4. Unclassifiable | 3 | 18 | 3 | 1 |

*Note.* The Kuhlman/Teta measure was given prior to the two experimental sequences, and the Ring Measure was given after the two sequences.

Given that generosity itself varied linearly with cooperative behavior, it appears that the willingness to give to others is an important underlying determinant of the SVO typology.

Several limitations of this study are noteworthy. First, the generalizability and relevance of two-person PD data to nonexperimental settings have been questioned (Messick & Brewer, 1983; Pruitt & Kimmel, 1977). Thus, it would be desirable to extend the present results to more naturalistic contexts and to multiperson dilemmas. Second, the two experiments in this study occurred only 1 month apart. To enhance confidence that participants had really learned the lessons of tit-for-tat, one should use a longer time interval. Third, it would be desirable to examine the influence of other programmed strategies (besides tit-for-tat) on longer term increases in cooperative behavior, as well as the influence of other payoff matrices and personality variables on longer term change. This would reveal still more about how lasting cooperation is effected. Fourth, these questions could be studied from the perspective of theories of personality or ego development (Loevinger, 1997). What dynamic and intrapsychic processes accompany or enable an individual's developmental shift from selfish to prosocial values and strategies?

## Conclusion

The PD efficiently captures the functional form of a wide variety of human transactions, in which (fragile) mutual cooperation provides the best mutual outcome. Thus, it is important to understand how people can learn, over time, to be more cooperative in situations of this type. The present findings suggest that sustained learning occurs when individuals encounter the firm but fair tit-for-tat strategy. When this happens, they move in the direction of more cooperative goals, stronger expectations of cooperation from others, and, thus, more cooperative behavior. Of course, some have more need of this lesson than others. It is encouraging that, in this research, those who needed the lesson the most were also the ones who paid the most heed.

## References


Axelrod, R. (1984). *The evolution of cooperation.* New York: Basic Books.

Axelrod, R. (1997). *The complexity of cooperation.* Princeton, NJ: Princeton University Press.

De Dreu, C., Giebels, E., & Van de Vliet, E. (1998). Social motives and trust in integrative negotiation: The disruptive effects of punitive capacity. *Journal of Applied Social Psychology, 83,* 408–422.

Fry, R., Betz, B., & Pruitt, G. (1996). The effects of resource shortage on de-escalation in a simulated price war. *International Journal of Conflict Management, 7,* 5–20.

Graziano, W. G., & Eisenberg, N. H. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 795–824). San Diego, CA: Academic Press.

Griffin, D., & Gonzalez, R. (1995). Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin, 118,* 430–439.

Harford, T., & Solomon, L. (1967). "Reformed sinner" and "lapsed saint" strategies in the prisoner's dilemma game. *Journal of Conflict Resolution, 11,* 104–109.

Heatherton, T., & Weinberger, J. (1994). Conceptual issues in assessing whether personality can change. In T. Heatherton & J. Weinberger (Eds.), *Can personality change?* (pp. 3–18). Washington, DC: American Psychological Association.

Judd, C., & Kenny, D. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5,* 602–619.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence.* New York: Wiley.

Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology, 46,* 183–207.

Komorita, S. S., Parks, C. D., & Hulbert, L. G. (1992). Reciprocity and the induction of cooperation in social dilemmas. *Journal of Personality and Social Psychology, 62,* 607–617.

Kuhlman, D. M., Brown, C., & Teta, P. (1992). Judgments of cooperation and defection in social dilemmas: The moderating role of judge's social orientation. In W. B. Liebrand, D. M. Messick, & H. A. Wilke (Eds.), *Social dilemmas: Theoretical issues and research findings* (pp. 111–132). Oxford, England: Pergamon.

Kuhlman, D. M., Camac, C., & Cunha, D. A. (1986). Individual differences in social orientation. In H. Wilke, D. Messick, & C. Rutte (Eds.), *Experimental social dilemmas* (pp. 47–62). New York: Verlag Peter Lang.

Kuhlman, D. M., & Marshallo, A. (1975). Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma. *Journal of Personality and Social Psychology, 32,* 922–931.

Laszlo, E., Artigiani, R., Combs, A., & Csanyi, V. (1996). *Changing visions: Human cognitive maps: Past, present, and future.* Westport, CT: Praeger.

Liebrand, W. B. (1984). The effect of social motives, communication and group size on behavior in a N-person multi-stage mixed-motive game. *European Journal of Social Psychology, 14,* 239–264.

Loevinger, J. (1997). Stages of personality development. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 198–208). San Diego, CA: Academic Press.

Messick, D. M., & Brewer, M. B. (1983). Solving social dilemmas: A review. In L. Wheeler & P. Shaver (Eds.), *Review of personality and social psychology* (Vol. 4, pp. 11–44). Beverly Hills, CA: Sage.

Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology, 4,* 1–25.

Oskamp, S. (1971). Effects of programmed strategies on cooperation in the prisoner's dilemma and other mixed-motive games. *Journal of Conflict Resolution, 15,* 225–259.

Pruitt, D. G. (1970). Motivational processes in the decomposed prisoner's dilemma game. *Journal of Personality and Social Psychology, 14,* 227–238.

Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology, 28,* 363–392.

Rusbult, C. E., & Van Lange, P. A. (1996). Interdependence processes. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 564–596). New York: Guilford Press.

Sober, E., & Wilson, D. (1998). *Unto others: The evolution and psychology of unselfish behavior.* Cambridge, MA: Harvard University Press.

Tognoli, J. (1975). Reciprocation of generosity and knowledge of game termination in the decomposed prisoner's dilemma game. *European Journal of Social Psychology, 5,* 297–312.

Van Lange, P. A. (1992). Rationality and morality in social dilemmas: The influence of social value orientations. In W. B. Liebrand, D. M. Messick, & H. A. Wilke (Eds.), *Social dilemmas: Theoretical issues and research findings* (pp. 133–146). Oxford, England: Pergamon.

Van Lange, P. A., Liebrand, W. B., Messick, D. M., & Wilke, H. A. (1992). Introduction and literature review. In W. B. Liebrand, D. M. Messick, & H. A. Wilke (Eds.), *Social dilemmas: Theoretical issues and research findings* (pp. 3–28). Oxford, England: Pergamon.

Yamagishi, T. (1992). Group size and the provision of a sanctioning system in a social dilemma. In W. B. Liebrand, D. M. Messick, & H. A. Wilke (Eds.), *Social dilemmas: Theoretical issues and research findings* (pp. 267–287). Oxford, England: Pergamon.


Received June 5, 1998
Revision received February 26, 1999
Accepted May 10, 1999 ■